

1652

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/471,349

DATE: 07/19/2000
 TIME: 15:40:27

Input Set : A:\PTO.txt
 Output Set: N:\CRF3\07192000\I471349.raw

ENTERED

```

4 <110> APPLICANT: Sahni, Girish
5   Kumar, Rajesh
6   Roy, Chaiti
7   Rajagopal, Kammara
8   Nihalani, Deepak
9   Sundaram, Vasudha
10  Yadav, Mahavir
12 <120> TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
13   PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
14   PREPARATION OF SAID PROTEINS
16 <130> FILE REFERENCE: 07064/009001
18 <140> CURRENT APPLICATION NUMBER: US 09/471,349
19 <141> CURRENT FILING DATE: 1999-12-23
21 <150> PRIOR APPLICATION NUMBER: IN 3825/DEL/98
22 <151> PRIOR FILING DATE: 1998-12-24
24 <160> NUMBER OF SEQ ID NOS: 24
26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1245
30 <212> TYPE: DNA
31 <213> ORGANISM: Streptococcus equisimilis
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (1)...(1242)
37 <400> SEQUENCE: 1
38 att gct gga cct gag tgg ctg cta gac cgt cca tct gtc aac aac agc      48
39 ile ala gly pro glu trp leu leu asp arg pro ser val asn asn ser
40 1      5      10      15
42 caa tta gtt gtt agc gtt gct ggt act gtt gag ggg acg aat caa gac      96
43 gln leu val val ser val ala gly thr val glu gly thr asn gln asp
44      20      25      30
46 att agt ctt aaa ttt ttt gaa atc gat cta aca tca cga cct gct cat      144
47 ile ser leu lys phe phe glu ile asp leu thr ser arg pro ala his
48      35      40      45
50 gga gga aag aca gag caa ggc tta agt cca aaa tca aaa cca ttt gct      192
51 gly gly lys thr glu gln gly leu ser pro lys ser lys pro phe ala
52      50      55      60
54 act gat agt ggc gcg atg tca cat aaa ctt gag aaa gct gac tta cta      240
55 thr asp ser gly ala met ser his lys leu glu lys ala asp leu leu
56      65      70      75      80
58 aag gct att caa gaa caa ttg atc gct aac gtc cac agt aac gac gac      288
59 lys ala ile gln glu gln leu ile ala asn val his ser asn asp asp
60      85      90      95
62 tac ttt gag gtc att gat ttt gca agc gat gca acc att act gat cga      336
63 tyr phe glu val ile asp phe ala ser asp ala thr ile thr asp arg
64      100      105      110
66 aac ggc aag gtc tac ttt gct gac aaa gat ggt tcg gta acc ttg ccg      384

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/471,349 DATE: 07/19/2000
 TIME: 15:40:27

Input Set : A:\PTO.txt
 Output Set: N:\CRF3\07192000\I471349.raw

67	Asn	Gly	Lys	Val	Tyr	Phe	Ala	Asp	Lys	Asp	Gly	Ser	Val	Thr	Leu	Pro	
68			115					120					125				
70	acc	caa	cct	gtc	caa	gaa	ttt	ttg	cta	agc	gga	cat	gtg	cgc	gtt	aga	432
71	Thr	Gln	Pro	Val	Gln	Glu	Phe	Leu	Leu	Ser	Gly	His	Val	Arg	Val	Arg	
72			130					135					140				
74	cca	tat	aaa	gaa	aaa	cca	ata	caa	aac	caa	gcg	aaa	tct	gtt	gat	gtg	480
75	Pro	Tyr	Lys	Glu	Lys	Pro	Ile	Gln	Asn	Gln	Ala	Lys	Ser	Val	Asp	Val	
76	145						150				155				160		
78	gaa	tat	act	gta	cag	ttt	act	ccc	tta	aac	cct	gat	gac	gat	ttc	aga	528
79	Glu	Tyr	Thr	Val	Gln	Phe	Thr	Pro	Leu	Asn	Pro	Asp	Asp	Asp	Phe	Arg	
80					165				170					175			
82	cca	ggt	ctc	aaa	gat	act	aag	cta	ttg	aaa	aca	cta	gct	atc	ggt	gac	576
83	Pro	Gly	Leu	Lys	Asp	Thr	Lys	Leu	Lys	Thr	Leu	Ala	Ile	Gly	Asp		
84			180					185					190				
86	acc	atc	aca	tct	caa	gaa	tta	cta	gct	caa	gca	caa	agc	att	tta	aac	624
87	Thr	Ile	Thr	Ser	Gln	Glu	Leu	Leu	Ala	Gln	Ala	Gln	Ser	Ile	Leu	Asn	
88			195				200					205					
90	aaa	aac	cac	cca	ggc	tat	acg	att	tat	gaa	cgt	gac	tcc	tca	atc	gtc	672
91	Lys	Asn	His	Pro	Gly	Tyr	Thr	Ile	Tyr	Glu	Arg	Asp	Ser	Ser	Ile	Val	
92		210				215					220						
94	act	cat	gac	aat	gac	att	ttc	cgt	acg	att	tta	cca	atg	gat	caa	gag	720
95	Thr	His	Asp	Asn	Asp	Ile	Phe	Arg	Thr	Ile	Leu	Pro	Met	Asp	Gln	Glu	
96	225			230				235				240					
98	ttt	act	tac	cgt	gtt	aaa	aat	cgg	gaa	caa	gct	tat	agg	atc	aat	aaa	768
99	Phe	Thr	Tyr	Arg	Val	Lys	Asn	Arg	Glu	Gln	Ala	Tyr	Arg	Ile	Asn	Lys	
100			245					250				255					
102	aaa	tct	ggt	ctg	aat	gaa	gaa	ata	aac	aac	act	gac	ctg	atc	tct	gag	816
103	Lys	Ser	Gly	Leu	Asn	Glu	Glu	Ile	Asn	Asn	Thr	Asp	Leu	Ile	Ser	Glu	
104			260					265				270					
106	aaa	tat	tac	gtc	ctt	aaa	aaa	ggg	gaa	aag	ccg	tat	gat	ccc	ttt	gat	864
107	Lys	Tyr	Tyr	Val	Leu	Lys	Lys	Gly	Glu	Lys	Pro	Tyr	Asp	Pro	Phe	Asp	
108			275				280					285					
110	cgc	agt	cac	ttg	aaa	ctg	ttc	acc	atc	aaa	tac	gtt	gat	gtc	gat	acc	912
111	Arg	Ser	His	Leu	Lys	Leu	Phe	Thr	Ile	Lys	Tyr	Val	Asp	Val	Asp	Thr	
112		290				295					300						
114	aac	gaa	ttg	cta	aaa	agt	gag	cag	ctc	tta	aca	gct	agc	gaa	cgt	aac	960
115	Asn	Glu	Leu	Leu	Lys	Ser	Glu	Gln	Leu	Leu	Thr	Ala	Ser	Glu	Arg	Asn	
116	305				310				315				320				
118	tta	gac	ttc	aga	gat	tta	tac	gat	cct	cgt	gat	aag	gct	aaa	cta	ctc	1008
119	Leu	Asp	Phe	Arg	Asp	Leu	Tyr	Asp	Pro	Arg	Asp	Lys	Ala	Lys	Leu	Leu	
120			325					330				335					
122	tac	aac	aat	ctc	gat	gct	ttt	ggt	att	atg	gac	tat	acc	tta	act	gga	1056
123	Tyr	Asn	Asn	Leu	Asp	Ala	Phe	Gly	Ile	Met	Asp	Tyr	Thr	Leu	Thr	Gly	
124			340					345				350					
126	aaa	gta	gag	gat	aat	cac	gat	gac	acc	aac	cgt	atc	ata	acc	gtt	tat	1104
127	Lys	Val	Glu	Asp	Asn	His	Asp	Asp	Thr	Asn	Arg	Ile	Ile	Thr	Val	Tyr	
128			355				360					365					
130	atg	ggc	aag	cga	ccc	gaa	gga	gag	aat	gct	agc	tat	cat	tta	gcc	tat	1152
131	Met	Gly	Lys	Arg	Pro	Glu	Gly	Glu	Asn	Ala	Ser	Tyr	His	Leu	Ala	Tyr	

RAW SEQUENCE LISTING DATE: 07/19/2000
 PATENT APPLICATION: US/09/471,349 TIME: 15:40:27

Input Set : A:\PTO.txt
 Output Set: N:\CRF3\07192000\I471349.raw

```

132      370      375      380
134 gat aaa gat cgt tat acc gaa gaa gaa cga gaa gtt tac agc tac ctg      1200
135 Asp Lys Asp Arg Tyr Thr Glu Glu Glu Arg Glu Val Tyr Ser Tyr Leu
136 385      390      395      400
138 cgt tat aca ggg aca cct ata cct gat aac cct aac gac aaa      1242
139 Arg Tyr Thr Gly Thr Pro Ile Pro Asp Asn Pro Asn Asp Lys
140      405      410
142 taa      1245
144 <210> SEQ ID NO: 2
145 <211> LENGTH: 414
146 <212> TYPE: PRT
147 <213> ORGANISM: Streptococcus equisimilis
149 <400> SEQUENCE: 2
150 Ile Ala Gly Pro Glu Trp Leu Leu Asp Arg Pro Ser Val Asn Asn Ser
151 1      5      10      15
152 Gln Leu Val Val Ser Val Ala Gly Thr Val Glu Gly Thr Asn Gln Asp
153      20      25      30
154 Ile Ser Leu Lys Phe Phe Glu Ile Asp Leu Thr Ser Arg Pro Ala His
155      35      40      45
156 Gly Gly Lys Thr Glu Gln Gly Leu Ser Pro Lys Ser Lys Pro Phe Ala
157 50      55      60
158 Thr Asp Ser Gly Ala Met Ser His Lys Leu Glu Lys Ala Asp Leu Leu
159 65      70      75      80
160 Lys Ala Ile Gln Glu Gln Leu Ile Ala Asn Val His Ser Asn Asp Asp
161      85      90      95
162 Tyr Phe Glu Val Ile Asp Phe Ala Ser Asp Ala Thr Ile Thr Asp Arg
163      100      105      110
164 Asn Gly Lys Val Tyr Phe Ala Asp Lys Asp Gly Ser Val Thr Leu Pro
165      115      120      125
166 Thr Gln Pro Val Gln Glu Phe Leu Leu Ser Gly His Val Arg Val Arg
167      130      135      140
168 Pro Tyr Lys Glu Lys Pro Ile Gln Asn Gln Ala Lys Ser Val Asp Val
169 145      150      155      160
170 Glu Tyr Thr Val Gln Phe Thr Pro Leu Asn Pro Asp Asp Asp Phe Arg
171      165      170      175
172 Pro Gly Leu Lys Asp Thr Lys Leu Leu Lys Thr Leu Ala Ile Gly Asp
173      180      185      190
174 Thr Ile Thr Ser Gln Glu Leu Leu Ala Gln Ala Gln Ser Ile Leu Asn
175      195      200      205
176 Lys Asn His Pro Gly Tyr Thr Ile Tyr Glu Arg Asp Ser Ser Ile Val
177      210      215      220
178 Thr His Asp Asn Asp Ile Phe Arg Thr Ile Leu Pro Met Asp Gln Glu
179 225      230      235      240
180 Phe Thr Tyr Arg Val Lys Asn Arg Glu Gln Ala Tyr Arg Ile Asn Lys
181      245      250      255
182 Lys Ser Gly Leu Asn Glu Glu Ile Asn Asn Thr Asp Leu Ile Ser Glu
183      260      265      270
184 Lys Tyr Tyr Val Leu Lys Lys Gly Glu Lys Pro Tyr Asp Pro Phe Asp
185      275      280      285

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/471,349

DATE: 07/19/2000
 TIME: 15:40:27

Input Set : A:\PTO.txt
 Output Set: N:\CRF3\07192000\I471349.raw

```

186 Arg Ser His Leu Lys Leu Phe Thr Ile Lys Tyr Val Asp Val Asp Thr
187      290      295      300
188 Asn Glu Leu Leu Lys Ser Glu Gln Leu Leu Thr Ala Ser Glu Arg Asn
189 305      310      315      320
190 Leu Asp Phe Arg Asp Leu Tyr Asp Pro Arg Asp Lys Ala Lys Leu Leu
191      325      330      335
192 Tyr Asn Asn Leu Asp Ala Phe Gly Ile Met Asp Tyr Thr Leu Thr Gly
193      340      345      350
194 Lys Val Glu Asp Asn His Asp Asp Thr Asn Arg Ile Ile Thr Val Tyr
195      355      360      365
196 Met Gly Lys Arg Pro Glu Gly Glu Asn Ala Ser Tyr His Leu Ala Tyr
197      370      375      380
198 Asp Lys Asp Arg Tyr Thr Glu Glu Arg Glu Val Tyr Ser Tyr Leu
199 385      390      395      400
200 Arg Tyr Thr Gly Thr Pro Ile Pro Asp Asn Pro Asn Asp Lys
201      405      410
203 <210> SEQ ID NO: 3
204 <211> LENGTH: 777
205 <212> TYPE: DNA
206 <213> ORGANISM: Homo sapiens
208 <220> FEATURE:
209 <221> NAME/KEY: CDS
210 <222> LOCATION: (1)...(777)
212 <400> SEQUENCE: 3
213 cag gct cag caa atg gtt cag ccc cag tcc ccg gtg gct gtc agt caa      48
214 Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
215 1      5      10      15
217 agc aag ccc ggt tgt tat gac aat gga aaa cac tat cag ata aat caa      96
218 Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile Asn Gln
219      20      25      30
221 cag tgg gag cgg acc tac cta ggt aat gtg ttg gtt tgt act tgt tat      144
222 Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val Cys Thr Cys Tyr
223      35      40      45
225 gga gga agc cga ggt ttt aac tgc gaa agt aaa cct gaa gct gaa gag      192
226 Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro Glu Ala Glu Glu
227      50      55      60
229 act tgc ttt gac aag tac act ggg aac act tac cga gtg ggt gac act      240
230 Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr
231 65      70      75      80
233 tat gag cgt cct aaa gac tcc atg atc tgg gac tgt acc tgc atc ggg      288
234 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly
235      85      90      95
237 gct ggg cga ggg aga ata agc tgt acc atc gca aac cgc tgc cat gaa      336
238 Ala Gly Arg Gly Arg Ile Ser Cys Thr Ile Ala Asn Arg Cys His Glu
239      100      105      110
241 ggg ggt cag tcc tac aag att ggt gac acc tgg agg aga cca cat gag      384
242 Gly Gly Gln Ser Tyr Lys Ile Gly Asp Thr Trp Arg Arg Pro His Glu
243      115      120      125
245 act ggt ggt tac atg tta gag tgt gtg tgt ctt ggt aat gga aaa gga      432

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/471,349

DATE: 07/19/2000

TIME: 15:40:27

Input Set : A:\PTO.txt

Output Set: N:\CRF3\07192000\I471349.raw

```

246 Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly Asn Gly Lys Gly
247      130      135      140
249 gaa tgg acc tgc aag ccc ata gct gag aag tgt ttt gat cat gct gct      480
250 Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe Asp His Ala Ala
251 145      150      155      160
253 ggg act tcc tat gtg gtc gga gaa acg tgg gag aag ccc tac caa ggc      528
254 Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys Pro Tyr Gln Gly
255      165      170      175
257 tgg atg atg gta gat tgt act tgc ctg gga gaa ggc agc gga cgc atc      576
258 Trp Met Met Val Asp Cys Thr Cys Leu Gly Glu Gly Ser Gly Arg Ile
259      180      185      190
261 act tgc act tct aga aat aga tgc aac gat cag gac aca agg aca tcc      624
262 Thr Cys Thr Ser Arg Asn Arg Cys Asn Asp Gln Asp Thr Arg Thr Ser
263      195      200      205
265 tat aga att gga gac acc tgg agc aag aag gat aat cga gga aac ctg      672
266 Tyr Arg Ile Gly Asp Thr Trp Ser Lys Lys Asp Asn Arg Gly Asn Leu
267      210      215      220
269 ctc cag tgc atc tgc aca ggc aac ggc cga gga gag tgg aag tgt gag      720
270 Leu Gln Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu Trp Lys Cys Glu
271 225      230      235      240
273 agg cac acc tct gtg cag acc aca tcg agc gga tct ggc ccc ttc acc      768
274 Arg His Thr Ser Val Gln Thr Thr Ser Ser Gly Ser Gly Pro Phe Thr
275      245      250      255
277 gat gtt cgt      777
278 Asp Val Arg
281 <210> SEQ ID NO: 4
282 <211> LENGTH: 259
283 <212> TYPE: PRT
284 <213> ORGANISM: Homo sapiens
286 <400> SEQUENCE: 4
287 Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
288 1      5      10      15
289 Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile Asn Gln
290      20      25      30
291 Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val Cys Thr Cys Tyr
292      35      40      45
293 Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro Glu Ala Glu Glu
294      50      55      60
295 Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr
296 65      70      75      80
297 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly
298      85      90      95
299 Ala Gly Arg Gly Arg Ile Ser Cys Thr Ile Ala Asn Arg Cys His Glu
300      100      105      110
301 Gly Gly Gln Ser Tyr Lys Ile Gly Asp Thr Trp Arg Arg Pro His Glu
302      115      120      125
303 Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly Asn Gly Lys Gly
304      130      135      140
305 Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe Asp His Ala Ala

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/471,349

DATE: 07/19/2000

TIME: 15:40:28

Input Set : A:\PTO.txt

Output Set: N:\CRF3\07192000\I471349.raw

L:698 M:283 W: Missing Blank Line separator, <210> field identifier